UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/051,462	01/22/2002	Hans-Gunter Hirsch	4114-3	6967

23117 7590 12/08/2005

NIXON & VANDERHYE, PC
901 NORTH GLEBE ROAD, 11TH FLOOR
ARLINGTON, VA 22203

EXAMINER	
LERNER, MARTIN	
ART UNIT	**PAPER NUMBER**
2654	

DATE MAILED: 12/08/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/051,462	HIRSCH ET AL.	
	Examiner	Art Unit	
	Martin Lerner	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *10 November 2005*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1 to 26* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1 to 26* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *10 November 2005* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Drawings

1. The drawings are objected to because they are informal. Applicants' Replacement Sheets for Figures 1 to 3 submitted 10 November 2005 are acceptable as formal drawings, but contain lines down the right side of the sheets due to facsimile transmission. Applicants are requested to resubmit Figures 1 to 3 to eliminate the obscuring lines.

2. Corrected drawing sheets in compliance with 37 CFR 1.121(d) are required in reply to the Office action to avoid abandonment of the application. Any amended replacement-drawing sheet should include all of the figures appearing on the immediate prior version of the sheet, even if only one figure is being amended. The figure or figure number of an amended drawing should not be labeled as "amended." If a drawing figure is to be canceled, the appropriate figure must be removed from the replacement sheet, and where necessary, the remaining figures must be renumbered and appropriate changes made to the brief description of the several views of the drawings for consistency. Additional replacement sheets may be necessary to show the renumbering of the remaining figures. Each drawing sheet submitted after the filing date of an application must be labeled in the top margin as either "Replacement Sheet" or "New Sheet" pursuant to 37 CFR 1.121(d). If the examiner does not accept the

changes, the applicant will be notified and informed of any required corrective action in the next Office action. The objection to the drawings will not be held in abeyance.

Claim Rejections - 35 USC § 112

3. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

4. Claims 21 to 26 are rejected under 35 U.S.C. 112, first paragraph, as failing to comply with the written description requirement. The claims contains subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventors, at the time the application was filed, had possession of the claimed invention.

The limitation of a method that "does not include an averaging over a complete utterance" is new matter because it is not described in the originally-filed Specification. Applicants' Specification doesn't say averaging is not used. There is no express disclosure for the limitation of "does not include an averaging over a complete utterance" in Applicants' originally-filed Specification. Moreover, Applicants' Equation on Page 15, Lines 26, is equivalent to an averaging over a complete utterance. Page 15, Lines 34 to 37 clearly states that the factor 1/T serves for an averaging or normalization to the length of the sequence. Thus, the limitation of "does not include an averaging over a complete utterance" represents new matter.

Claim Rejections - 35 USC § 102

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

6. Claims 1, 3 to 13, 15 to 17, and 19 are rejected under 35 U.S.C. 102(b) as being anticipated by *Takagi ('057)*.

Regarding independent claims 1 and 15, *Takagi ('057)* discloses a speech recognition method, program code, and device compensating for background noise, comprising:

"providing a set of reference speech spectra" – reference pattern 3 is words or sentences of speech of a standard speaker that have been analyzed (column 5, lines 1 to 5: Figure 1);

"determining the reference speech spectra which correspond to the distorted short-term speech spectra" – spectra X(t), Y(t), V(t), and W(t) are "short-term" because they represent a cepstral vector at a discrete time t (column 4, lines 44 to 46 and column 4, lines 59 to 60); noise conditions of additive noise and channel distortion of recognized input speech and those of the reference pattern are matched (column 8, lines 34 to 46); a reference pattern is analyzed and matched to feature vectors of the input speech (column 5, lines 5 to 18); implicitly, feature vectors represent "short-term

speech spectra" because feature vectors correspond to one frame of speech, which is the shortest time period for speech analysis;

"estimating a frequency response taking into account both the distorted short-term speech spectra and the corresponding reference speech spectra" – spectral transforming portion 4 transforms the time sequence X(t) of the feature vectors of the input speech and the time sequence Y(t) of the feature vectors of the reference pattern into time sequences V(t) and W(t) of spectra; cepstra are transformed into spectra (column 5, lines 18 to 30: Figure 1); *Takagi ('057)*'s filters A_v and A_w are equivalent to Applicants' $[H(f)]^2$ frequency response (Compare Applicants' Equation on Page 1, Line 37 to *Takagi ('057)*'s Equations (3) and (4));

"compensating the distorted short-term speech spectra based on the estimated frequency response" – compensating portion 6 matches additive noise and channel distortion of the input speech with those of the reference pattern corresponding to Equations (11) and (13); compensation is performed by multiplying one of the reference pattern and the input speech by a predetermined channel distortion so that the average value of the speech pattern becomes equal to that of the input speech (column 8, lines 4 to 21: Figure 1); here, multiplying the input speech by a predetermined channel distortion provides for "compensating the distorted short-term speech spectra".

Regarding independent claims 17 and 19, *Takagi ('057)* further discloses a database for storing reference speech spectra because reference patterns 3 are

implicitly stored in a database element, as illustrated (Figure 1); additionally, a processor implicitly performs the method steps of the flowchart (Figure 1).

Regarding claim 3, *Takagi ('057)* discloses compensating speech as a spectrum of the input speech and a reference pattern ("in the spectral domain").

Regarding claim 4, *Takagi ('057)* discloses that spectra of additive noise B_ω and channel distortion A_ω of a reference pattern are known (column 5, line 62 to column 6, line 12).

Regarding claims 5 and 7, *Takagi ('057)* discloses that additive noise and channel distortion of input speech is matched to those of the reference pattern (column 6, lines 13 to 17); matching involves finding a closest reference pattern to input speech.

Regarding claim 6, *Takagi ('057)* discloses stored reference patterns 3 for speech recognition (column 5, lines 1 to 5: Figure 1); implicitly, reference patterns are known in the art as "models".

Regarding claims 8 and 13, *Takagi ('057)* discloses compensating a reference pattern by taking an average of input speech for regions of additive noise and channel distortion during preliminary matching 2 (column 6, lines 22 to 57: Figure 1).

Regarding claim 9, *Takagi ('057)* discloses matching input speech and reference patterns by a matching error (column 6, lines 8 to 12: Figure 1); a matching error represents a difference between input speech and a reference pattern.

Regarding claim 10, *Takagi ('057)* discloses average vector calculating portion 5 calculates the average vector of the time sequences of the spectra of the input speech (column 9, lines 3 to 8: Figure 1).

Regarding claims 11 and 12, *Takagi ('057)* discloses using average values of spectra of input speech and reference patterns (column 6, lines 13 to 17: Figure 1); an average is calculated by summing over previous samples K_Ω and K_Φ (column 6, lines 22 to 57); averaging over a number of past samples is equivalent to "smoothing".

Regarding claim 16, *Takagi ('057)* discloses a procedure described by a flowchart (Figure 1), which is implicitly performed on a digital signal processor, with a recording medium storing the instructions of the procedure.

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

8. Claims 2, 14, and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable over *Takagi ('057)* in view of *Takahashi.*

Concerning independent claim 14, *Takagi ('057)* discloses all the limitations, but does not expressly provide for "obtaining distorted speech spectra and analyzing the distorted speech spectra by means of a speech/nonspeech decision to filter out the distorted speech spectra that do not contain speech." In fact, however, *Takagi ('057)*

discloses storing predetermined speech regions and noise regions of reference patterns (column 5, lines 6 to 9), and using average values of speech and noise regions of the input speech (column 6, lines 13 to 17). Thus, while *Takagi ('057)* does not expressly disclose a speech/nonspeech decision filter to filter out distorted speech spectra that do not contain speech, implicitly, there must be a speech/nonspeech detector to decide which regions are speech regions and which regions are noise regions.

Those skilled in the art know that a voice activity detector (VAD) ("a speech/nonspeech decision filter") is a common element for making speech/nonspeech decisions for a variety of purposes in speech processing. Specifically, *Takahashi* teaches noise suppression for removing noise from voice, where a voice/nonvoice discriminator 32 judges whether a voice signal separated into frames is voice or non-voice. The objective is to estimate a noise spectrum during silent periods so as to subtract a noise spectrum from a distorted speech spectrum and thereby correct a distorted speech spectrum to eliminate noise (Column 7, Line 38 to Column 8, Line 11: Figure 4) It would have been obvious to one having ordinary skill in the art to analyze distorted speech with a speech/nonspeech decision as taught by *Takahashi* in the method of removing noise during speech recognition of *Takagi ('057)* for the purpose of estimating a noise spectrum during silent periods so that noise may be eliminated.

Concerning claim 2, similar considerations apply.

Concerning claim 18, *Takahashi* discloses first spectrum memory 36a and second spectrum memory 36b for temporarily storing prior frames of speech spectra

(column 7, lines 51 to 61), which are equivalent to "a buffer", a common expedient implicit in speech processing.

9. Claim 20 is rejected under 35 U.S.C. 103(a) as being unpatentable over *Takagi ('057)* in view of *Brown et al.*

Takagi ('057) discloses all of the limitations, omitting only "a distributed speech recognition system" having "a network server with central speech recognition means." However, distributed speech recognition with a client/server architecture and central speech recognition on a server are commonly known because more computationally intensive speech recognition activities may be performed on a server to minimize the computational requirements of a client. Specifically, *Brown et al.* teaches an acoustic speech recognizer system and method, where a phone browser 12 connects to speech recognition server 34. (Column 2, Line 23 to Column 3, Line 8: Figures 1 and 2) *Brown et al.* states an advantage of a speech recognizer system that has a barge-in detector discriminating between speech and noise, and does not need a push-to-talk command. (Column 1, Lines 35 to 56) It would have been obvious to one having ordinary skill in the art to incorporate a speech recognition apparatus of *Takagi ('057)* into a distributed speech recognition system with a central speech recognition server as suggested by *Brown et al.* for the purpose of eliminating a need for a push-to-talk button.

Response to Arguments

10. Applicant's arguments filed 10 November 2005 have been fully considered but they are not persuasive.

Firstly, Applicants argue that their independent claims set forth short-term speech spectra, and do not rely upon an averaging mechanism over a complete utterance. Applicants point to their Specification, Page 5, Lines 22 to 25, as defining "short-term" as denoting a period of time corresponding to a typical frame length. Applicants contend that the rejection takes an incorrect interpretation of the phrase "frequency response", which has a clear technical meaning, and should not be interpreted as simply an amplitude for each speech frequency. These arguments are not persuasive.

Even assuming Applicants' definition of short-term speech spectra, *Takagi ('057)* still equivalently discloses estimating a frequency response taking into account short-term speech spectra. Applicants' Specification, Page 5, Lines 22 to 25, defines "short-term" as denoting a period of time corresponding to a typical frame length, but Page 5, Lines 25 to 37, then goes on to say that each of the short-term spectra is sequentially processed, and may contain all of the short-term spectra within an utterance. The passage also states that there are some cases where not every short-term spectra may be taken into account to estimate a frequency response, as when only every second short-term spectra is sufficient to base an estimation, but only in an extreme case is it suggested that that a few or a single short-term spectra is input to estimate a frequency response.

Takagi ('057) performs averaging over short-term time sequences of spectra X(t), Y(t), V(t), and W(t). Each of spectra X(t), Y(t), V(t), and W(t) are "short-term" because they represent a cepstral vector at a discrete time t. (Column 4, Lines 44 to 46 and Column 4, Lines 59 to 60) A vector representing a discrete time t corresponds to a frame of speech, and thus, meets Applicants' definition of "short-term" spectra. Time sequences of spectra X(t), Y(t), V(t), and W(t) do not represent more than one frame for *Takagi ('057)*.

Applicants' attempt to distinguish their invention by maintaining that *Takagi ('057)* discloses an averaging mechanism while their independent claims do no rely upon an averaging mechanism is not persuasive. Applicants' Equation on Page 15, Line 24, clearly represents an averaging mechanism. Applicants' Specification, Page 15, Lines 34 to 37, states that the factor 1/T represents an averaging or normalization over the length of the sequence to take into account the number of short-term speech spectra in an utterance. Thus, Applicants' argument that their independent claims should be distinguished as based upon short-term speech spectra without the necessity of performing an averaging over a complete utterance is contrary to what is disclosed by their Specification. Both Applicants and *Takagi ('057)* begin with short-term speech spectra or time sequences for a discrete time t of individual speech frames, and then average a sequence of frame to compensate for distortion. Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

Newly added claims 21 to 26 set forth new matter because Applicants' originally-filed Specification does not disclose, either expressly or impliedly, that the method of estimating a frequency response does not include an averaging over a complete utterance. The limitation implies that averaging is not utilized at all, and that implication is contrary to Applicants' Specification. If Applicants had claimed that a frequency response is estimated based upon averaging of every second short-term speech spectra, then such a claimed limitation would not involve new matter.

Moreover, assuming that the term "frequency response" has a clear technical meaning as Applicants' H(f), then *Takagi ('057)* still discloses corresponding frequency responses A_v and A_w. Applicants' $[H(f)]^2$ represents a frequency response of a filter placed onto a distorted speech spectra that is equivalent to *Takagi ('057)*'s filter A_v. *Takagi ('057)* places filter A_v in front of distorted speech signal V(hat)(t) in the presence of additive noise B_v so as to obtain a clean speech signal V(t). This can be compared to Applicants' Equation on Page 1, Line 37, where Y(t, f) corresponds to V(t), $[H(f)]^2$ corresponds to A_v, S(t, f) corresponds to V(hat)(t), and N(f) corresponds to B_v. Thus, Applicants recognize that their frequency response $[H(f)]^2$ is equal to *Takagi ('057)*'s A_v.

Secondly, Applicants argue that *Takagi ('057)* does not take into account both the distorted short-term speech spectra and the corresponding reference speech spectra, as indicated by the Equation on Page 15, line 24 of the Specification. Applicants maintain that *Takagi ('057)* takes into account only the distorted speech as given by Equation (9), or only the reference speech as given by Equation (10), but not both, as stated by independent claim 1. This position is not persuasive.

Takagi ('057) takes into account both the distorted short-term speech spectra and the corresponding reference speech spectra, as claimed, even if these are not shown in one equation. Applicants apparently contend that because *Takagi ('057)* does not disclose their Equation on Page 15, Line 24, incorporating terms for both short-term spectra and reference spectra, that *Takagi ('057)* cannot anticipate independent claim 1. However, *Takagi ('057)* does clearly disclose taking into account both the short-term distorted speech spectra and the reference spectra to compensate for distortion. Column 9, Lines 9 to 19, states that the compensating portion 6 compensates "at least one of the time sequences of the spectra of the input speech and the reference pattern". *Takagi ('057)* discloses taking into account of compensation for both the distorted speech given by Equation (9) and the reference spectra of Equation (10) because "at least one" is compensated, implying that both may be compensated. Indeed, *Takagi ('057)*'s Equation (12), at Column 7, Lines 55 to 56, is transformed so as to include terms for both distorted speech spectra V(hat)(t) and corresponding reference spectra W(hat)(t) as elements of a frequency response. Furthermore, "taking into account" both the distorted short-term speech spectra and the reference spectra does not necessarily imply that both are positively compensated, but only keeping track of which is compensated and which is not compensated, if either, but not both, may be compensated.

Therefore, the rejections of claims 1, 3 to 13, 15 to 17, and 19 under 35 U.S.C. 102(b) as being anticipated by *Takagi ('057)*, of claims 2, 14, and 18 under 35 U.S.C. 103(a) as being unpatentable over *Takagi ('057)* in view of *Takahashi*, of claim

20 under 35 U.S.C. 103(a) as being unpatentable over *Takagi ('057)* in view of *Brown et al.*, and of claims 21 to 26 under 35 U.S.C. 112, 1st ¶, as failing to comply with the written description requirement, are proper.

Conclusion

11. Applicants' amendment necessitated the new ground of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicants are reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

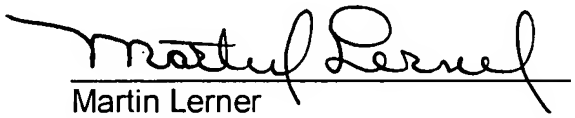
A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Martin Lerner whose telephone number is (571) 272-7608. The examiner can normally be reached on 8:30 AM to 6:00 PM Monday to Thursday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

ML
12/5/05



Martin Lerner
Examiner
Group Art Unit 2654